



... for a brighter future



U.S. Department
of Energy



A U.S. Department of Energy laboratory
managed by The University of Chicago

X-Ray Software Initiative

Kenneth Evans, Jr.

APS Engineering Support Division

May 4, 2006

2006 Users Meetings

Argonne, Illinois

Topics

- History and Motivation
- Example Software Frameworks
 - DANSE / Pyre
 - GumTree
 - ISAW
 - JAS3
 - OAG Apps
 - XRAYs
- Criteria for a Data Analysis Software Project
- Eclipse
 - IDE
 - RCP
- Visualization

The APS Has a Stated Commitment

- From the 5-year plan:

"Attaining our goal of maximizing the scientific productivity and impact will demand not only changes in beamline configuration but also improved detectors and software for data reduction and analysis. . . . Providing **robust data reduction, data analysis, modeling and simulation software** to our users is an important enabler to increase the volume of high-impact APS results. This program has **very high priority**, and will be pursued simultaneously with the highest priority beam line upgrades."

Committee on Software Framework Options

- Committee formed July 2005
 - Ken Evans, Guy Jennings, Tim Mooney
- Charge
 - Investigate possibilities for providing scientific software for users of the APS
 - Investigate DANSE
 - *Interact with the DANSE / Pyre developers*
 - *Implement a test case within their framework*
 - *Determine a set of criteria for producing facility supported analysis software at the APS that incorporates high end computing*
- This was accomplished
- The committee has expanded to include
 - Tom Worlton (IPNS), Gabrielle Long, Brian Toby

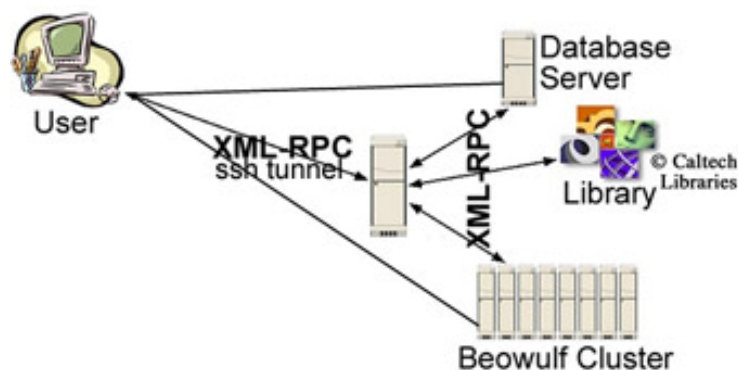
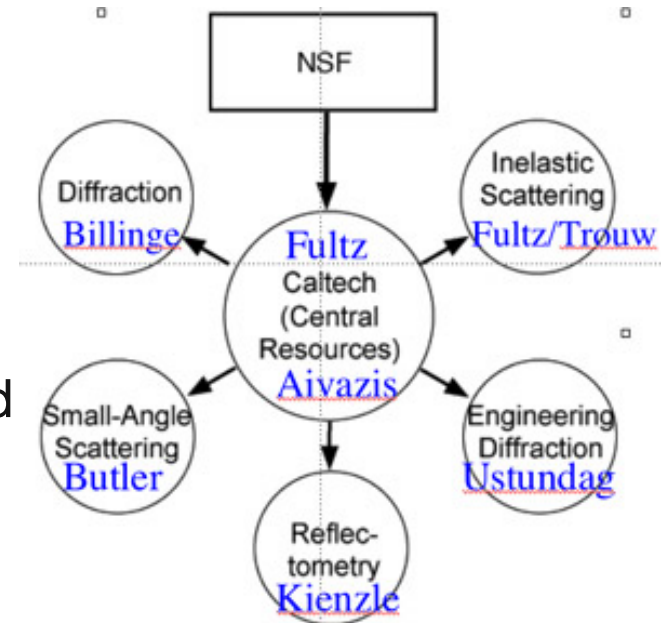
DANSE

- Data Analysis for Neutron Scattering Experiments
- Prompted by the Spallation Neutron Source (SNS)
- Recently funded by NSF after two years of proposals
 - \$12M for 5 years
 - Includes a computing cluster at Caltech
- Based on the Pyre Framework developed at Caltech
 - Python and C++
 - Approximately 10 years old
 - Extensive framework, but poorly documented
 - Under the control of one person



DANSE

- Principal scientific investigator is Brent Fultz, Caltech
 - 5 Sub-projects
 - Leaders are typically at universities
- Software engineering and technology is lead by Michael Aivazis
 - Director of Research, Center for Advanced Computing Research, Caltech
 - Principal Pyre author



From: DANSE Wiki

Study High Points

- Learned DANSE as a framework does not exist
- Pyre was downloaded and built
 - Windows and Linux – Build was successful
 - Solaris – Did not build
 - *Required almost a month of interaction with Michael Aivazis*
 - *Most of the debugging done here*
 - *Resulted in changes to C++ and Python modules*
 - *Are still problems with the build system*
- A test case was implemented
 - A Channel Access Probe
 - *Implements a GUI in Python*
 - *Uses “legacy” code*
 - *The C++ code was converted to Python using SWIG*
 - *Uses many Pyre features including interchangeable modules and persistence*

Pyre Evaluation

■ Advantages

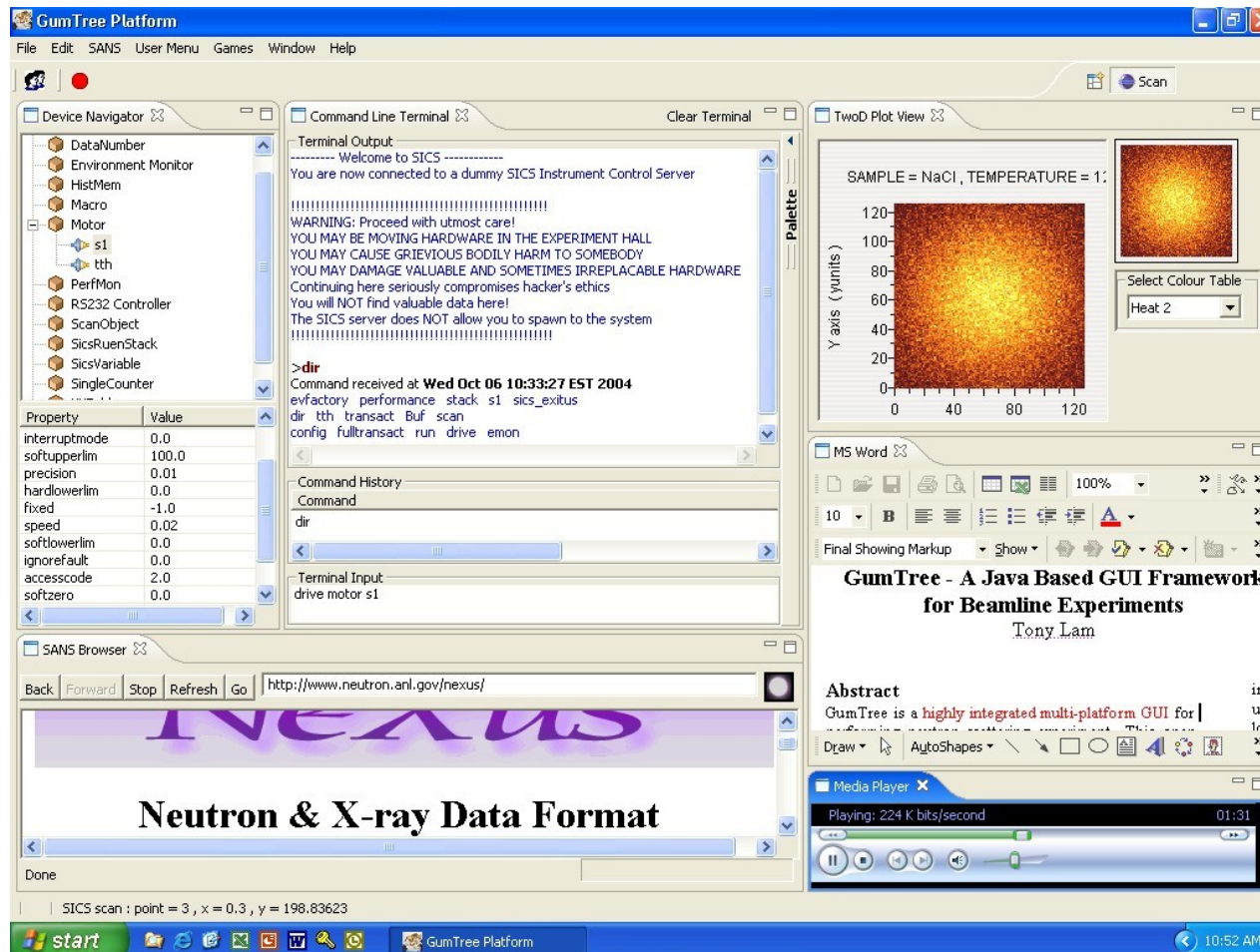
- Interchangeable Component model
- Extensive mechanisms for persistence
- Extensive error, debugging facility
- Message Passing Interface (MPI) implementation
- Other packages such as Nexus are already built in
- Multi-platform build model (but flawed)
- Python is an attractive level of complexity
- DANSE is using it

■ Disadvantages

- Documentation is severely lacking
- Relies on one individual – Michael Aivazis
- Is a sledge hammer for simple applications
- Is not designed with GUI applications in mind

Gum Tree

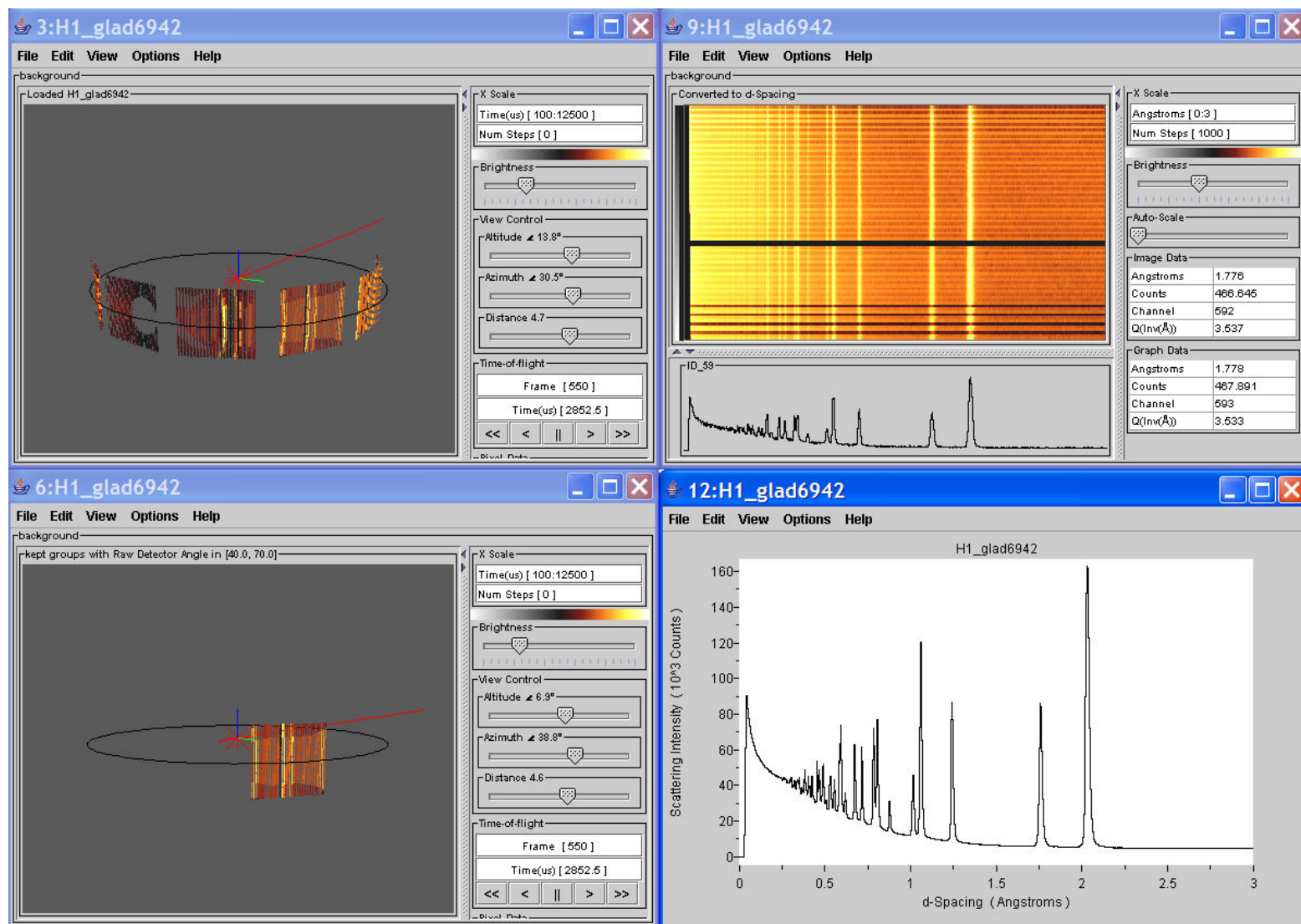
- GumTree is an Eclipse based GUI framework from ANSTO for running scientific experiments



From: Tony Lam, ICALEPCS Presentation, October 2004

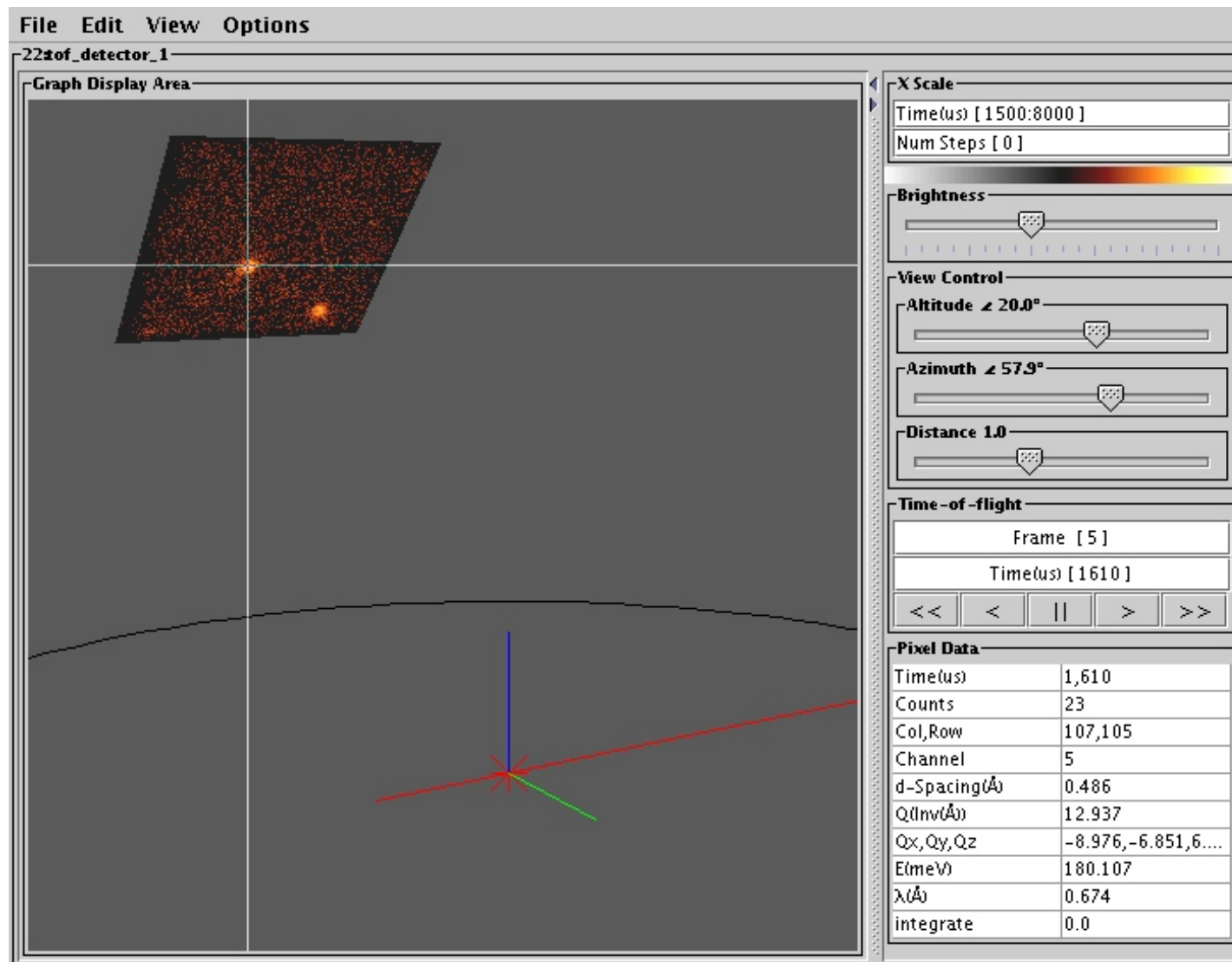
Current GumTree Status

- Workbench
- Tony Lam is the main developer
- Is currently in the process of redesign
- There is collaboration with
 - ESRF (Andy Goetz)
 - EPICS Control System Studio (CSS) (Mattias Clausen)
 - APS
- Three projects
 - GumTree RCP
 - *Integrated Scientific Workbench*
 - *Plug-in based*
 - Data Acquisition
 - Data Analysis
- Has promise



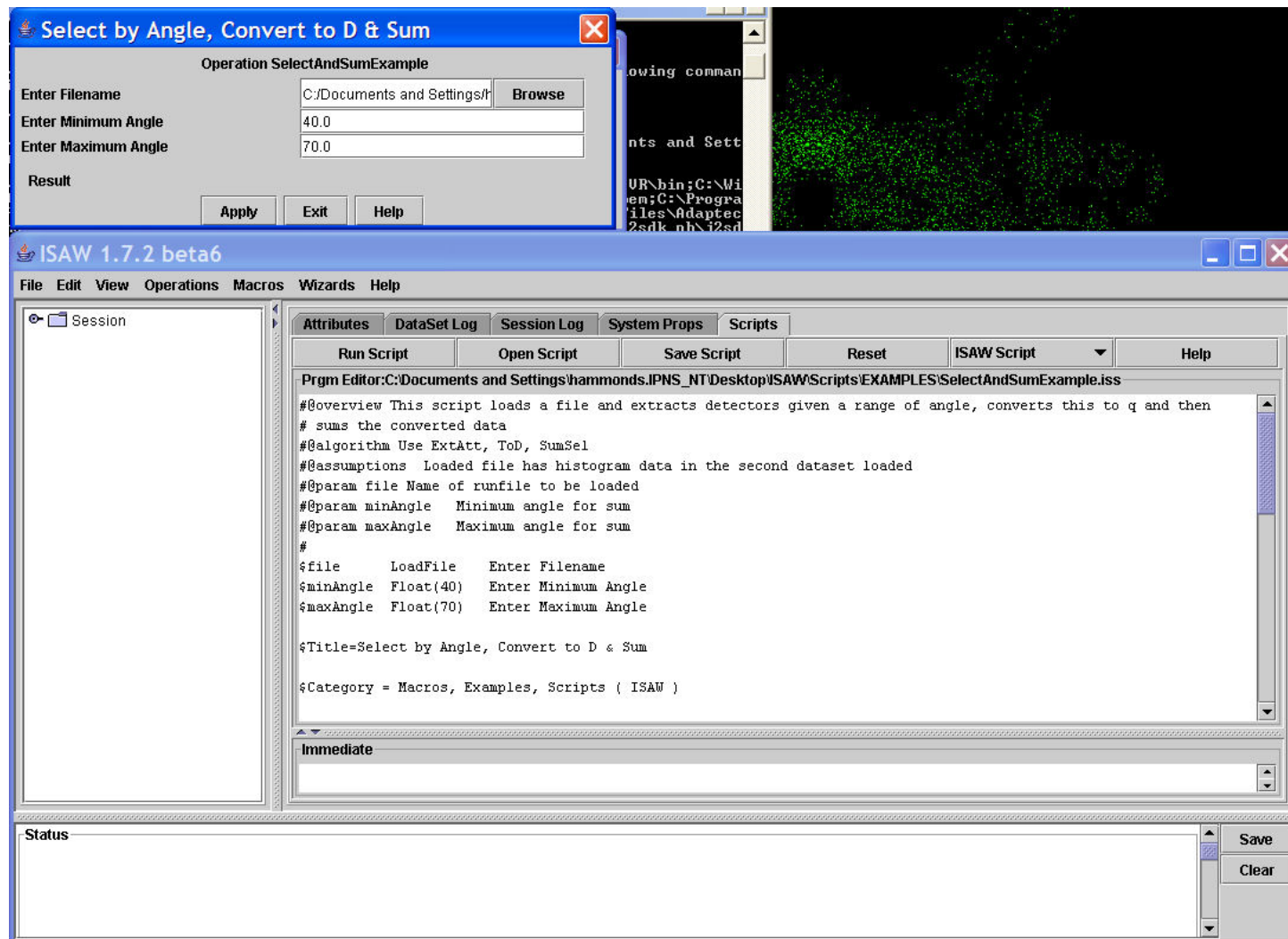
From: John Hammonds, IPNS

ISAW (3D View)



From: Dennis Mikkelsen, UW Stout State

ISAW (Scripts)



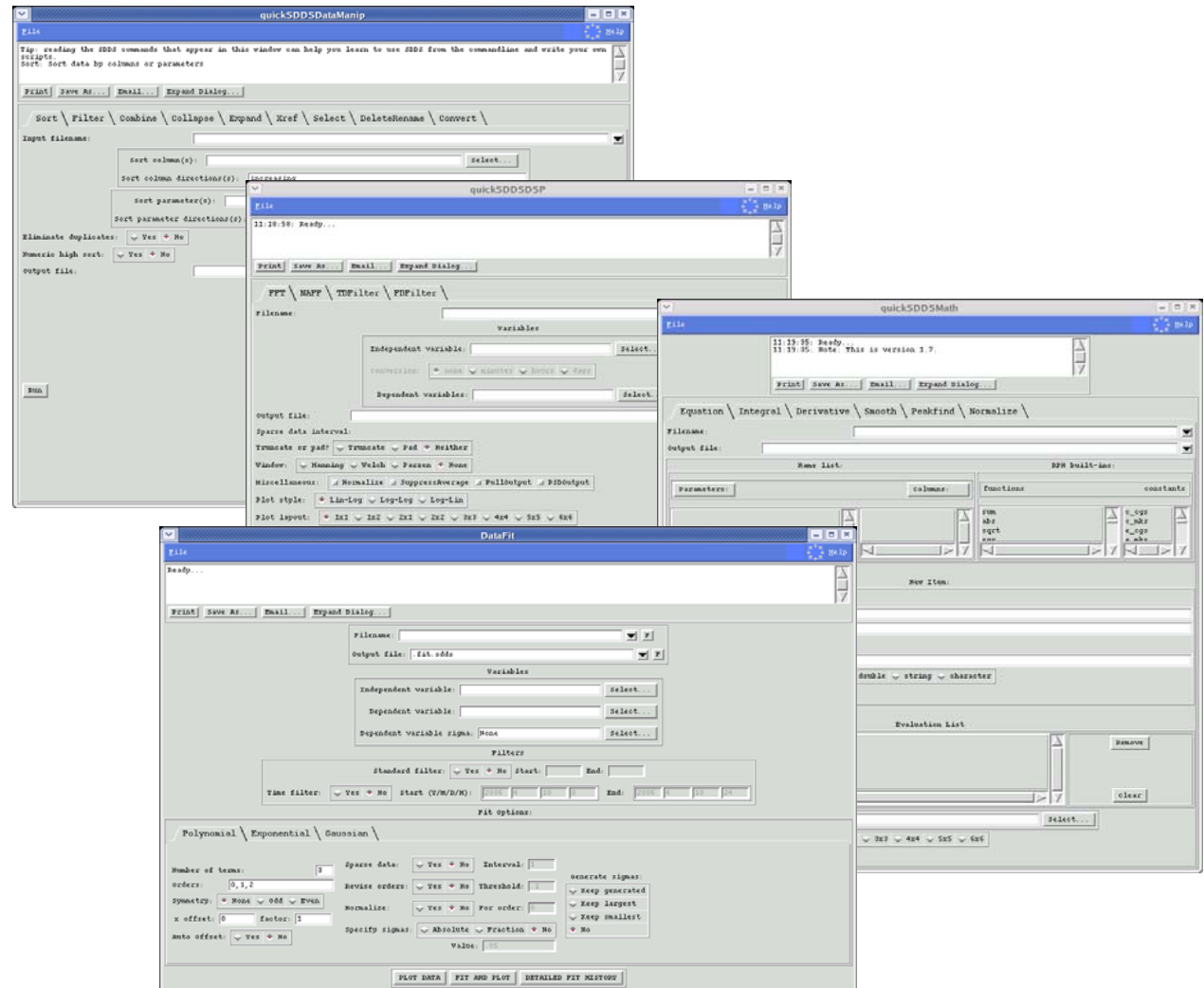
From: John Hammonds, IPNS

Current ISAW Status

- Workbench
- Stands for Integrated Spectral Analysis Workbench
- Started in 1999
 - John Hammonds, Dennis Mikkelsen, Tom Worlton
- 10 – 20 man years of effort
 - Above, Post Docs, Students (currently 10), Ruth Mikkelsen
- Collaborate with LANSCE
- Mikkelsons and their students are currently NSF funded
- Is a functioning tool
 - For IPNS data
 - Primarily time-of-flight data

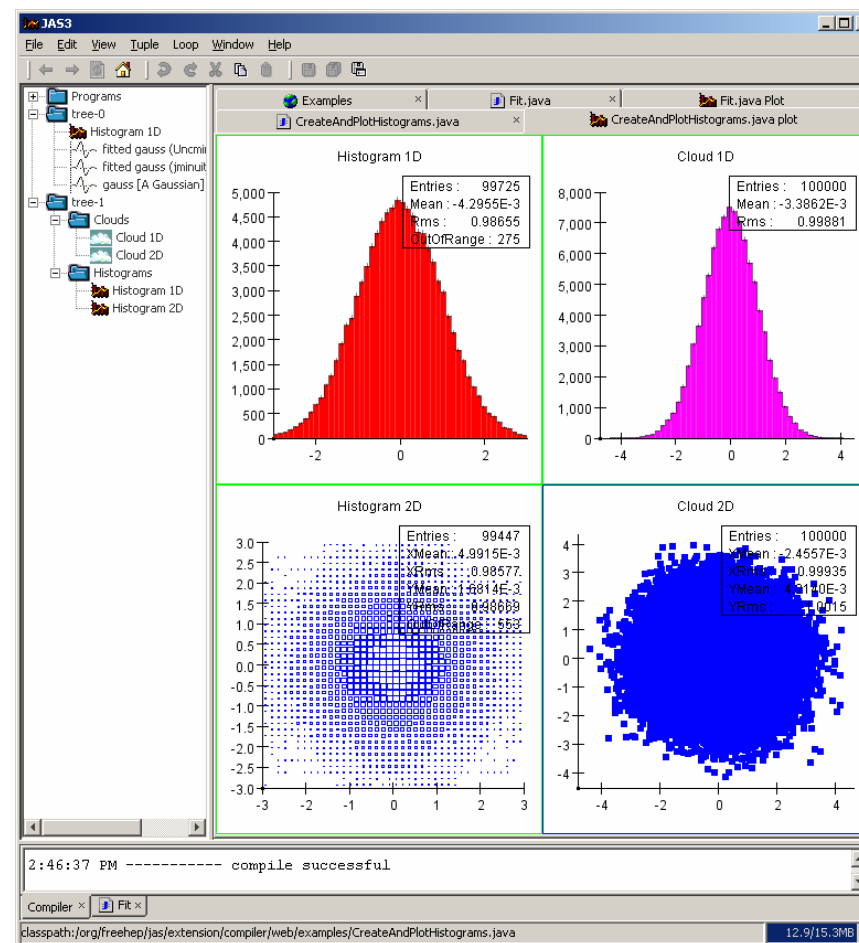
APS Operations & Analysis Group Tools

- Toolbox
- Based on Tcl/Tk
- Uses SDDS Toolkit



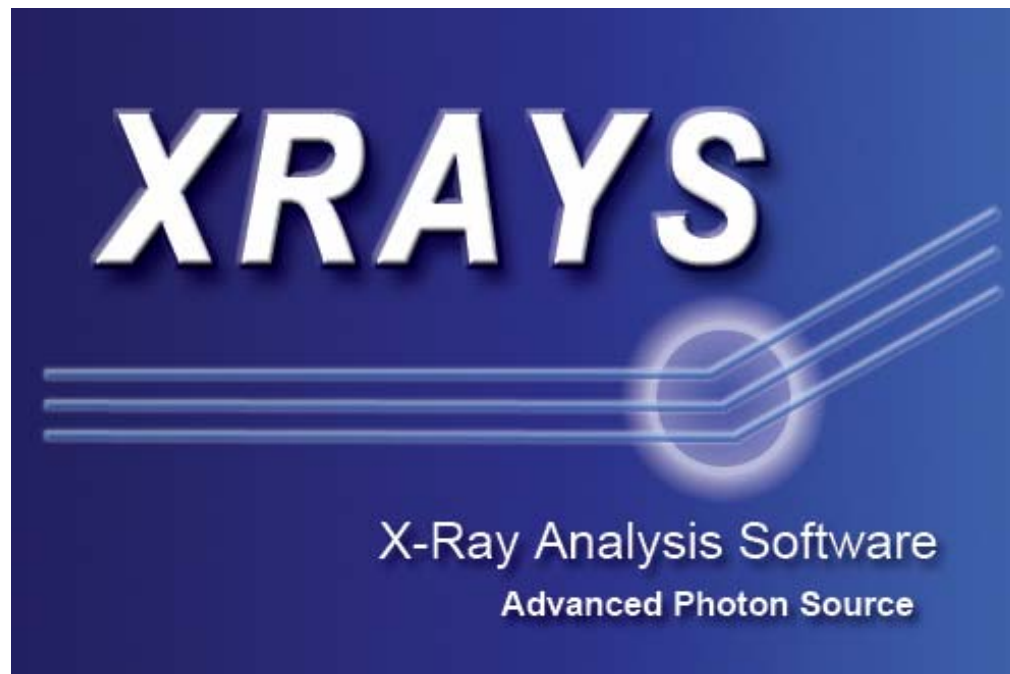
Java Analysis Studio (JAS3)

- Workbench
- Developed by / for the High-Energy physics community
- Plotting of 1d, 2d, 3d Histograms, XY plots, Scatter plots, *etc.*
- Open source
- Attractive plotting
- Fitting, other mathematical analysis
 - Primarily from CERN
- GUI for analysis tasks
- Built-in editor and compiler
- Highly modular structure
 - Uses plug-ins
- Documentation is lacking
- Does not do 2D array data well
 - i.e. CCD, Image plates



XRAYS

- X-Ray Analysis Software
- Currently in “Finding the Ballpark” stage
- Intended to be for the X-ray community what DANSE, GumTree, and ISAW are for the Neutron Scattering community
- Expect significant overlap with Neutron Scattering software



Criteria for a Data Analysis Software Project

- Needs to be a reasonably large project
 - To be better than what users are doing on their own
 - To provide continuity and eliminate dependence on key individuals
 - To do what meaningfully needs to be done
- Needs the approval of the users and community
 - Means it needs marketing and collaborative interaction
 - Community must see it as better than what they are doing now
- Needs to collaborate with the neutron-scattering community
 - There is much overlap
 - Resources can be shared
- Needs to provide high-performance computing
 - Not clear

Criteria for a Data Analysis Software Project

- Needs to concentrate on the higher-level process
 - Acquire – Filter – Analyze – View – Store
- Needs to be flexible and extensible
 - User's needs are diverse and varied
 - There are already many existing programs and approaches
 - *These need to be incorporated, not replaced*
 - Appropriate for novice as well as sophisticated users
- Needs to be cross platform
 - At least Linux and Windows, probably Solaris, hopefully OS X
- Needs to be language agnostic
 - Many analysis tools are FORTRAN, C, C++, and other



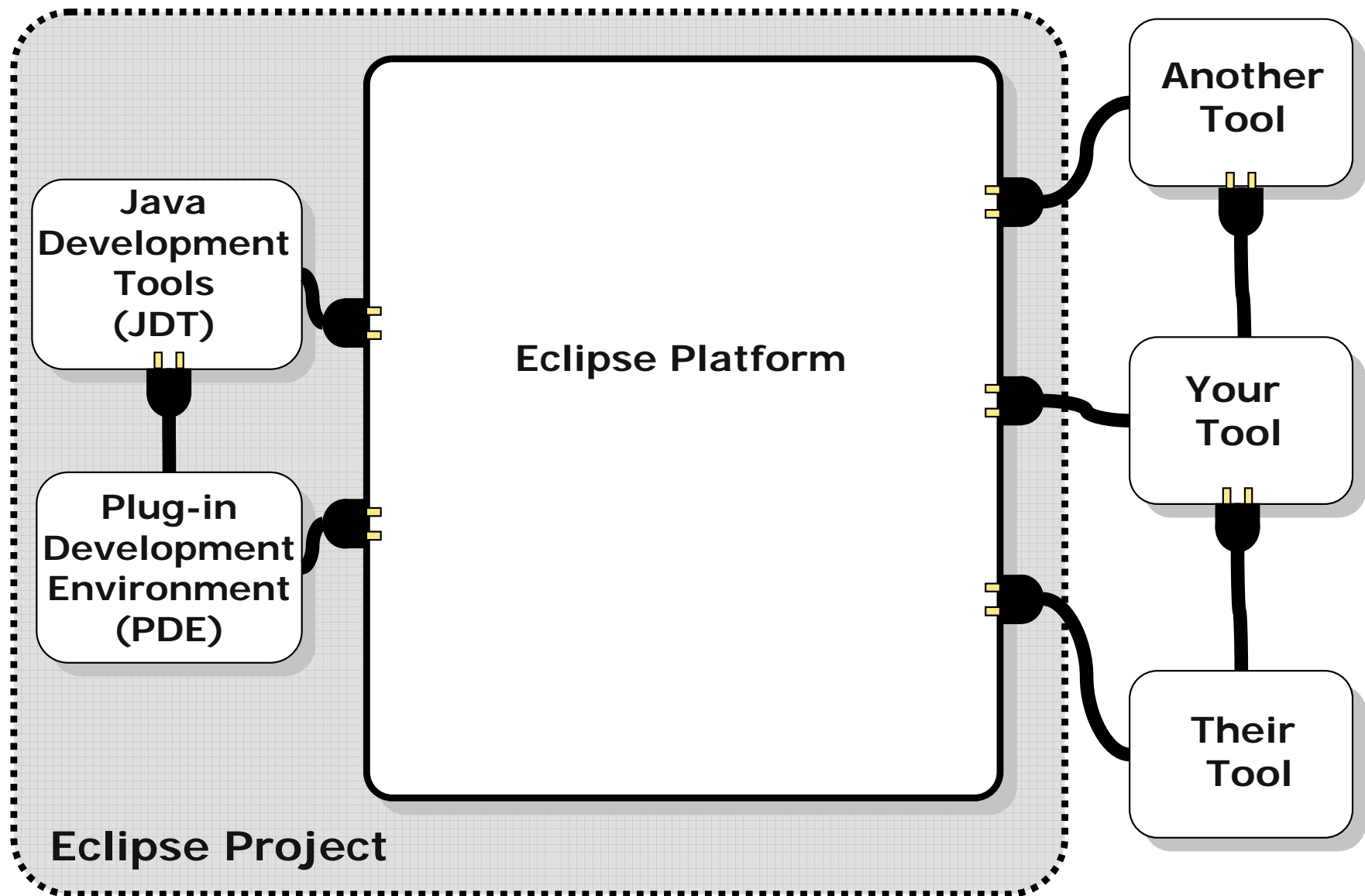
Criteria for a Data Analysis Software Project

- Needs to be easy to use
 - Scientists are more interested in Science than Computing
- Needs a common data format
 - Not easy
 - NeXus ?, SDDS ?
 - Perhaps data converters are the best pragmatic solution
- Needs to start with a smaller scope that is expandable
 - The current resources are limited
 - Easier than getting an entire, grand project like DANSE off the ground
 - Probably limited to the APS scope at first
- Needs to address two different types of requirements
 - Workbench
 - Toolbox

Eclipse

- Eclipse is an Open Source community
- It was started in 2001 by IBM
 - IBM donated a lot of research
 - Controlled the early development, but later relinquished control
- Out of the box it looks like a Java IDE
- It is really a Plug-in manager
 - That comes with Java Development plug-ins.
 - You can take these out and put your own (and/or others) in

Eclipse Overview



Modified From: Tony Lam, ICALEPCS Presentation, October 2004

Eclipse Foundation Membership

- Strategic Developers (13 as of Jan 2006)
 - At least 8 developers assigned full time to developing Eclipse
 - Contribution up to \$250K
- Strategic Consumers (4)
 - Contribution up to \$500K
 - Can reduce the dues by contributing 1-2 developers
- Three other tiers

- Bottom line
 - \$\$\$ and Developers (currently > 150 full time)
 - Compare to Pyre

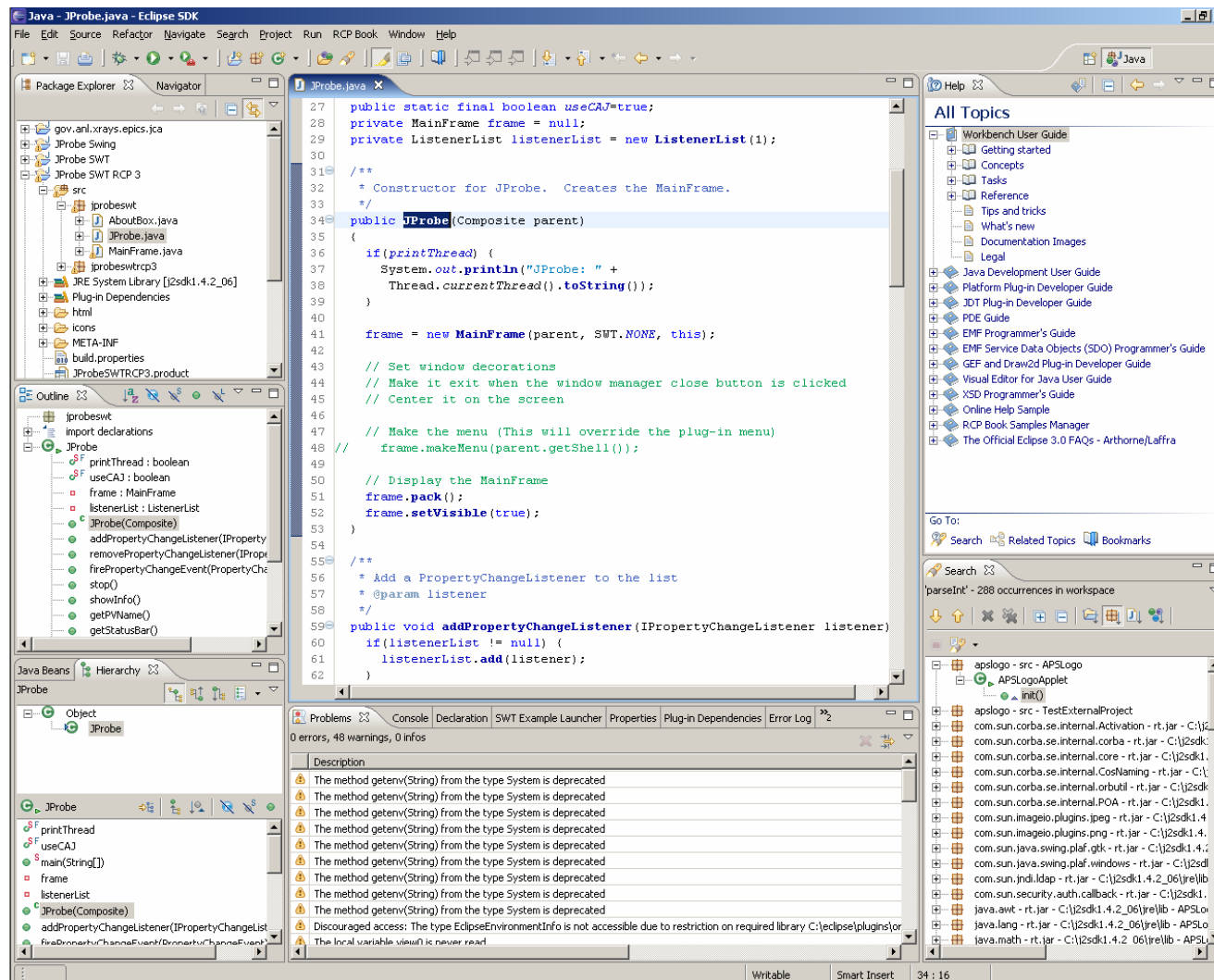
Eclipse Consortium Strategic Members



* Strategic Consumer

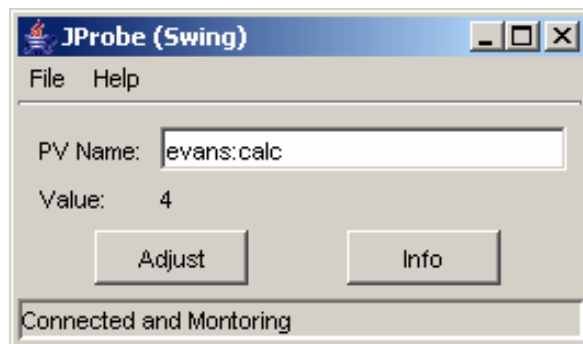
Eclipse as a Java IDE

- For developers, not users

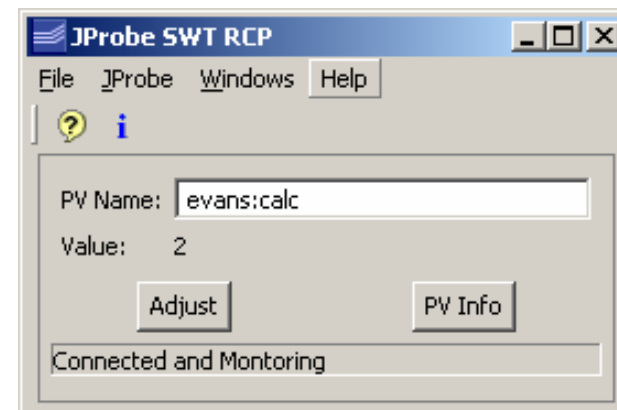


Eclipse As a Rich Client Platform

- Looks like an application, not an IDE
- Inherits a lot of functionality
 - Persistence (Properties and Preferences)
 - Help
 - Featured About dialog (like Eclipse's)
 - Splash screen
 - Dockable windows, and much more ...

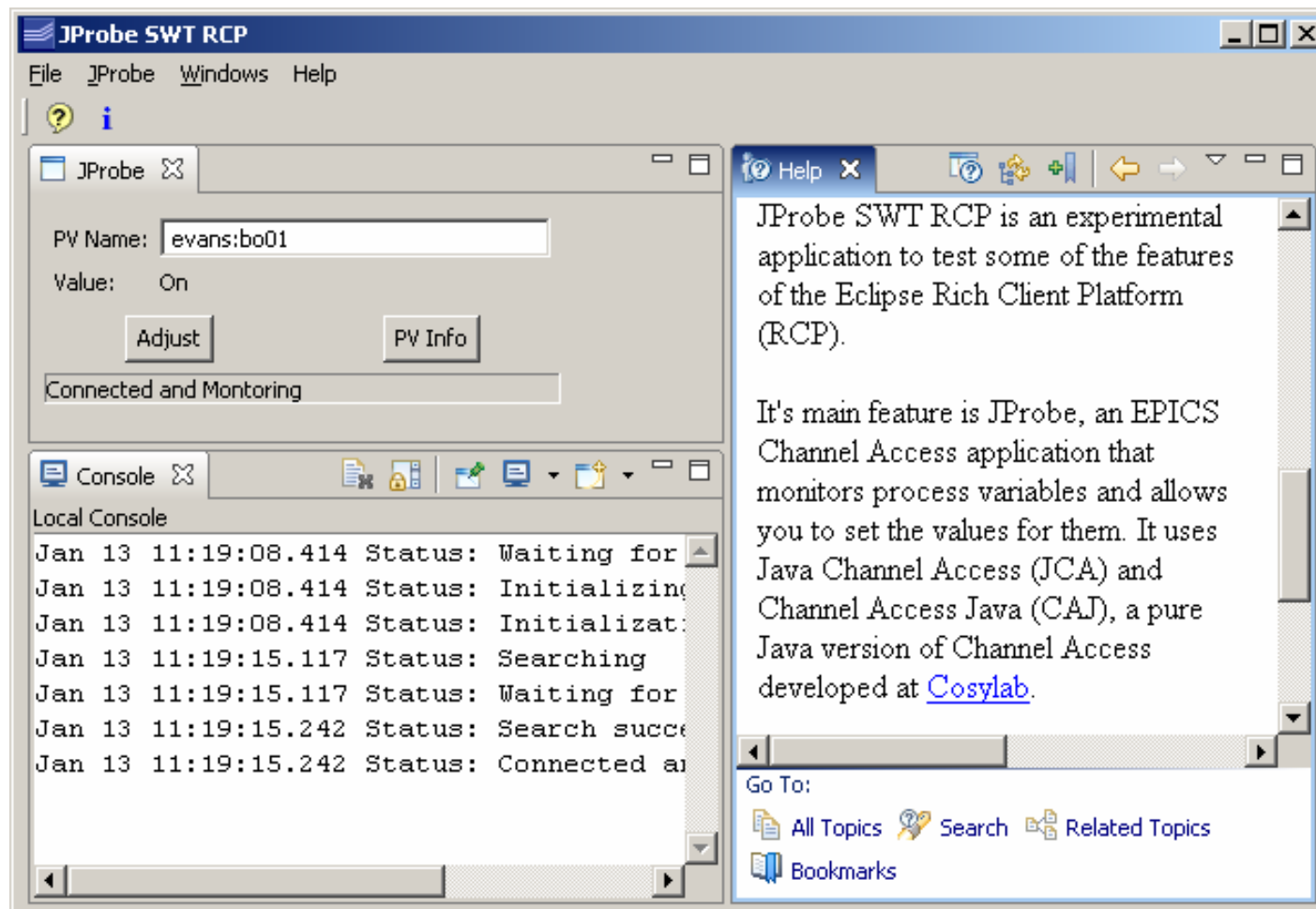


Java Application



RCP Application

Easily Made More Elaborate



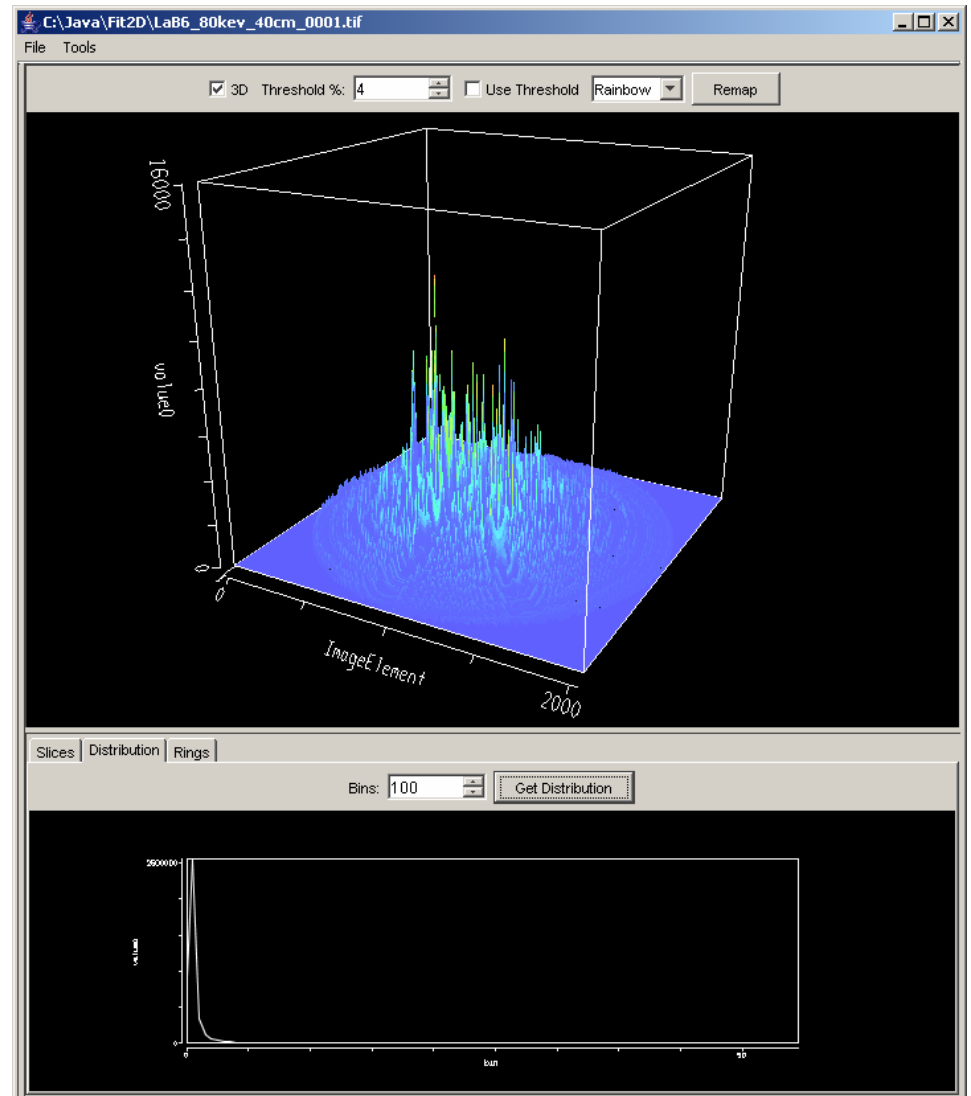
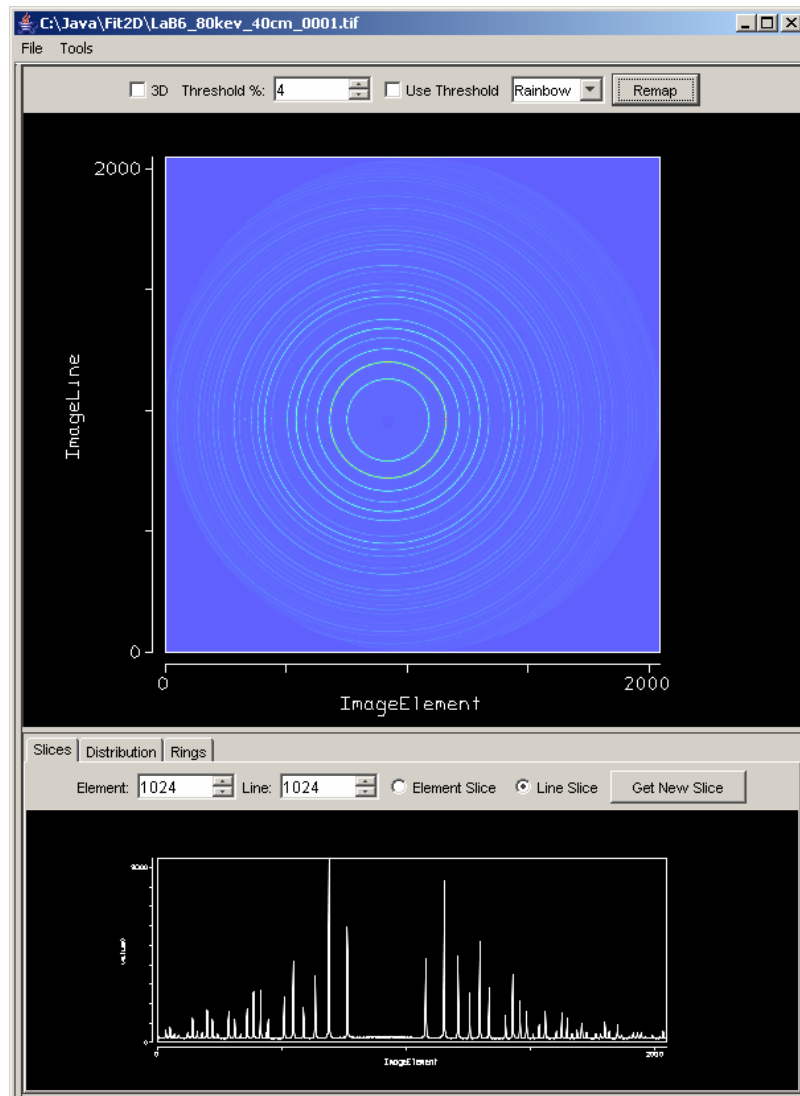
Bottom Line

- Is a very powerful and extensible IDE and Framework
- Is Open Source
- Has a community
- Is supported by most of the industry
- Has a large number of developers (>150)
- Has significant financial backing
- Are many 3rd-party Plug-ins, both free and commercial
- Is continuing to expand and improve rapidly
- Is free
- Downsides
 - Is a continually changing, moving target

Graphics and Visualization

- An essential part of a software analysis project
- Require many man years to develop
- We have looked at two packages
- FreeHEP (used by JAS3)
 - Based at SLAC (4 full-time people, 10 - 40 man years)
 - Attractive graphics
 - Documentation is lacking
 - Somewhat orthogonal to X-ray analysis needs
- VisAD
 - University of Wisconsin Space Science and Engineering Center
 - *And others*
 - Extensive package (10 years or more old)
 - Much documentation (but not enough)
 - Flexible and powerful

Prototype Image Analysis Tool using VisAD Graphics



Thank You

*This has been an
APS Controls Presentation*

Thank You

*This has been an
APS Controls Presentation*